Polarimetric second-harmonic generation microscopy of the hierarchical structure of collagen in stage I-III non-small cell lung carcinoma

AHMAD GOLARAEI,[1,2,3] (iD) **LEILA B. MOSTAÇO-GUIDOLIN,**[4] **VAISHNAVI RAJA,**[5] **ROYA NAVAB,**[3] **TAO WANG,**[6] **SHINGO SAKASHITA,**[3] **KAZUHIRO YASUFUKU,**[3,7] **MING-SOUND TSAO,**[3] **BRIAN C. WILSON,**[3,8] **AND VIRGINIJUS BARZDA**[1,2,9,*]

[1]*Department of Physics, University of Toronto, 60 St. George St, Toronto, M5S 1A7, Canada*
[2]*Department of Chemical and Physical Sciences, University of Toronto Mississauga, 3359 Mississauga Rd North, Mississauga, L5L 1C6, Canada*
[3]*Princess Margaret Cancer Centre, University Health Network, 101 College St, Toronto, M5G 1L7, Canada*
[4]*Department of Systems and Computer Engineering, Faculty of Engineering and Design, Carleton University, 1125 Colonel By Drive, Ottawa, K1S 5B6, Canada*
[5]*Department of Chemistry, University of Western Ontario, 1151 Richmond St, London, N6A 3K7, Canada*
[6]*Department of Pathology and Molecular Medicine, Queen's University, 88 Stuart St, Kingston, K7L 3N6, Canada*
[7]*Toronto General Hospital, University Health Network, 200 Elizabeth St, Toronto, M5G 2C4, Canada*
[8]*Department of Medical Biophysics, 101 College St, Toronto, M5G 1L7, Canada*
[9]*Laser Research Centre, Faculty of Physics, Vilnius University, Vilnius, 10223, Lithuania*
*virgis.barzda@utoronto.ca

Abstract: Polarimetric second-harmonic generation (P-SHG) microscopy is used to quantify the structural alteration of collagen in stage-I,-II and -III non-small cell lung carcinoma (NSCLC) *ex vivo* tissue. The achiral and chiral molecular second-order susceptibility tensor components ratios (R and C, respectively), the degree of linear polarization (DLP) and the in-plane collagen fiber orientation (δ) were extracted. Further, texture analysis was performed on the SHG intensity, R, C, DLP and δ. The distributions of R, C, DLP and δ as well as the textural features of entropy, correlation and contrast show significant differences between normal and tumor tissues.

1. Introduction

Lung cancer is the second most commonly diagnosed cancer worldwide with 85% being classified as non-small cell lung carcinoma (NSCLC) [1]. The most important prognostic factor for NSCLC, is staging, which is defined by the anatomic extent of the disease. Staging is usually done by the TNM method based on tumour (T), lymph nodes (N) and metastases (M) factors, the combination of which determine the stages I-IV; the last mostly involves metastatic spread. Staging determines the treatment strategy. While the focus of therapy for patients with stage IIIB-IV NSCLC is palliation, the treatment intent for stages I-IIIA is curative. The five-year survival rate for patients with stage-I NSCLC is 70%, dropping to around 1% for stage IV. Hence, early diagnosis and accurate staging are critical to impact survival rate and quality of life [2].

Tumor progression comprises not only tumor cell proliferation but also changes in the tumor microenvironment [3–5] that affects tumor growth and metastatic potential [6]. Previous studies have highlighted the impact of tumor growth on the structure and composition of the extracellular matrix (ECM) [7–10]. Since collagen is the major structure protein in the ECM, studying its structural alterations during tumor development has been the focus of many studies. These

alterations include degradation of collagen in the basement membrane [11] and remodeling of fibrillar collagen (mainly collagen type-I) throughout the connective tissue [12–16]. Although the structural alteration of collagen fibrils in NSCLC has been investigated previously [15,17], the impact of tumor stage on collagen structure is not well studied.

Since fibrillar collagen has a non-centrosymmetric structure, it can be visualized by second-harmonic generation (SHG) microscopy [18–20]. SHG is a coherent nonlinear process wherein two photons with the same frequency interact with a non-centrosymmetric material and produce a photon of double the frequency. SHG as a contrast mechanism has multiple advantages. The second-harmonic excitation is confined to a diffraction-limited volume which enables optical sectioning and three-dimensional (3D) imaging. Further, because SHG originates intrinsically from the biological structures, no staining is necessary. In addition, SHG does not require absorption for signal generation, and therefore, sample photobleaching and phototoxicity are reduced compared to other methods such as multiphoton fluorescence microscopy [21]. Further, the use of near-infrared excitation wavelengths makes deep tissue imaging possible [22]. Finally, the intensity and polarization of the generated SHG signal depend on the sample structure and organization, so that, by using polarimetric SHG microscopy (P-SHG) quantitative information can be obtained about the biomaterial structure and organization. One such technique is linear polarization-in, polarization-out (PIPO) SHG, where a set of incoming linear polarization states is prepared, for each of which a set of outgoing linear SHG polarizations is measured [23]. From these data, the achiral and chiral molecular second-order susceptibility tensor components ratios ($R = \chi_{zzz}/\chi_{zxx}$ and $C = \chi_{xyz}/\chi_{zxx}$, respectively), the in-plane fibril orientation angle (δ) and the degree of linear polarization (DLP) can be extracted, all of which reflect the structural properties of collagen on the sub-micrometer and micrometer scale [24,25]. The distributions of R, δ and DLP over scanned areas in histological sections have been successfully used to differentiate normal and malignant tissues in lung [17], thyroid [26], breast [27] and pancreas [28].

The distributions of R, C, δ and DLP values over the scanned areas provide ensemble characteristics of the image. On the other hand, the relation between the neighboring pixels, termed textural features, can also provide information on the morphological variations of the parameters over each scanned area. The extraction of textural features is based on a gray-level co-occurrence matrix (GLCM) and uses the second-order statistics of the grayscale image histograms [29,30]. Textural analysis of SHG intensity of collagen was previously used to study the alteration of collagen in cancerous tissues [12–14,16] and other pathological conditions [31–33]. In those studies, the entropy, inverse difference moment (IDM), contrast, and correlation of SHG intensity (I_{SHG}) were reported as the most suitable textural parameters to evaluate collagen changes [34–36]. The same analysis can be applied to R- and C-ratios, as well as δ and DLP images to extract additional textural information.

Here, we combine PIPO SHG microscopy measurements with GLCM texture analysis to investigate the changes in tumor-affected collagen across Stage I-III NSCLC. The alteration of collagen ultrastructure in tumor tissue is detected through R-, C-ratio and DLP from PIPO SHG microscopy. In addition, it is demonstrated that the tumor impacts the orientation of the collagen fibers which can be observed in the δ-distribution from PIPO SHG microscopy. Further, the differences in collagen texture are detected via entropy (S), correlation (ρ), and contrast (ν) from the texture analysis of I_{SHG}, while texture analyses of R, C, δ and DLP are less sensitive to the alterations of collagen texture, showing only trends that are not statistically significant. The hierarchical multiscale characterization of collagen structure has the potential to be used as a complementary technique for NSCLC staging.

2. Materials and methods

2.1. Tissue sample preparation

Tissues were collected according to an institutionally-approved protocol (University Health Network, Toronto, Canada) from 9 patients with NSCLC who underwent complete surgical resection of their tumors: 3 patients in each of stages-I, -II and -III. The normal-appearing lung parenchyma was taken at least 5 cm away from the tumor from the same patients. All samples were handled as per standard clinical histology protocols, generating formalin-fixed specimens cut into 5 μm thick sections and mounted on glass slides. The sections were stained with hematoxylin and eosin (H&E) and imaged with a bright-field microscope scanner (Aperio Whole Slide Scanner, Leica Biosystems) for reference. The stained samples were used for determining the regions of interest (ROIs) and they also have been used for SHG imaging without further preparation. For each patient, slides of one tumor and one normal lung section were investigated and at least 7 ROIs were scanned in each slide, as identified by a lung pathologist (T. W., S. S. or M.-S. T.), yielding a total of 23, 35 and 27 areas in stage-I, II and III specimens, respectively. Also 65 areas from the 9 patients in normal-appearing lung tissues were scanned as the control samples. For the controls, collagen rich patches in non-neoplastic lung were selected. For tumoral slides, areas of abnormal stroma that is contiguous with the cancer mass was selected as cancer associated stroma. Each scan area was 110 μm × 110 μm.

2.2. PIPO SHG microscope setup

The PIPO SHG imaging was performed with a custom nonlinear laser-scanning microscope, as described elsewhere [23]. The microscope was coupled with an in-house built diode-pumped Yb-ion-doped potassium gadolinium tungstate (Yb:KGW) crystal-based oscillator providing ~ 430 fs pulses at 1028 nm peak wavelength and 14.3 MHz pulse repetition rate [37]. The microscope contains a polarization state generator (PSG) to control the incident polarization state of the laser beam and a polarization state analyzer (PSA) to measure the polarization state of the outgoing SHG signal from the sample. The PSG has a linear polarizer (Laser Components Inc.) to ensure that the incident laser beam is linearly polarized and a half-wave plate (Comar Optics) to rotate the linear polarization of incoming laser beam. The PSG is placed before the excitation objective (20×, 0.75 NA air objective (Carl Zeiss)). The PSA contains only a rotating linear polarizer and it is placed after the custom 0.85 NA collection objective. The SHG signal was collected in the forward direction and detected with a single-photon counting photomultiplier tube (Hamamatsu H7422P-40). A BG 39 Schott glass filter and a 510-520 nm band-pass interference filter (Edmund Optics) were used in front of the detector to separate SHG from the laser light. The PSG half-wave plate was rotated to nine different angles with equal increments from 0° to 180°, while the PSA was rotated to nine different evenly spaced angles from 0° to 180° for each PSG state to perform the PIPO imaging. To monitor any possible sample degradation during imaging, the samples were periodically scanned at the initial PSG and PSA settings after measurements with each set of the PSA states. The 110 μm × 110 μm ROIs were scanned with a 2 μs dwell time, and 50 to 100 frames of 128 × 128 pixels were summed to obtain an image.

2.3. PIPO SHG image analysis

Ultrastructural properties of collagen fibers in tissue can be explained by the second-order susceptibility tensor components ($\chi^{(2)}_{ijk}$), where *ijk* denotes the Cartesian molecular coordinate system, with the *z*-axis being the collagen fiber direction. The 3D organization of collagen fibers in the lab coordinate system (denoted by *IJK/XYZ*) can be defined by the out-of-plane angle of the collagen fibril (α) measured from the image plane located in the *XZ*-plane and the average in-plane orientation of the fiber (δ) with respect to laboratory *Z*-axis. The laser propagation is set along the laboratory *Y*-axis. The detected SHG intensity from collagen will depend on

molecular susceptibility components as well as δ and α angles. It also depends on the orientation of the incoming laser polarization, θ, and orientation of the outgoing SHG polarization, ϕ, both measured with respect to the Z-axis of the laboratory frame of reference. The relation between the SHG intensity (I_{SHG}), polarizer angle (θ), analyzer angle (ϕ), the second-order susceptibility ratios (R and C), and in-plane orientation of the collagen fibrils (δ) can be formulated as follow [25]

$$I_{SHG} \propto \left| \sin(\phi - \delta) \sin 2(\theta - \delta) + \cos(\phi - \delta) \sin^2(\theta - \delta) \right.$$
$$\left. + R \cos(\phi - \delta) \cos^2(\theta - \delta) + 2C \cos(\theta - \delta) \sin(\theta - \phi) \right|^2 \tag{1}$$

where

$$R = \frac{\chi_{zzz}}{\chi_{zxx}} \cos^2 \alpha + 3 \sin^2 \alpha \tag{2a}$$

$$C = \frac{\chi_{xyz}}{\chi_{zxx}} \sin \alpha \tag{2b}$$

are the achiral and chiral ratio of in-image-plane molecular susceptibility components, respectively. In this equation, it is assumed that the sample birefringence is negligible, which is valid for the 5-µm thick tissue sections. χ_{xxz} and χ_{zxx} were shown to be similar in Achilles tendon collagen at the fundamental wavelength of 1028 nm [25], therefore $\chi_{xxz}/\chi_{zxx} = 1$ is assumed [38]. A trust region reflective (TRR) fitting [39] with custom written software in MATLAB (MathWorks) was performed on every pixel of the PIPO set of images using Eq. (1) as a function of θ and ϕ, and parameters R, C and δ were extracted. From PIPO SHG microscopy data, the degree of linear polarization (*DLP*) of the outgoing SHG signal can also be obtained. For each pixel of the image, the average fiber orientation angle, δ, extracted from the fit was used to select the closest PSG angle to the fiber axis, θ_c, such that $|\theta_c - \delta| \leq 11.25°$, where $11.25°$ is half of an incremental step of PSG. Since the SHG intensity depends on the angle between the incoming polarization and the collagen fiber axis orientation, the closest PSG angle was selected to decrease the intensity noise. In order to increase the signal-to-noise ratio (SNR), two sets of PSA orientation angles were used to calculate *DLP*: DLP_1 was calculated by Eq. (3a) and DLP_2 was calculated by Eq. (3b) where the subscript of intensity, I, indicates the PSA orientation angles.

$$DLP_1 = \frac{\sqrt{(I_0 - I_{90})^2 + (I_{45} - I_{135})^2}}{I_0 + I_{90}} \tag{3a}$$

$$DLP_2 = \frac{\sqrt{(I_{22.5} - I_{112.5})^2 + (I_{67.5} - I_{157.5})^2}}{I_{22.5} + I_{112.5}} \tag{3b}$$

The DLP_1 and DLP_2 were averaged to obtain the *DLP* for each pixel of the image. The *DLP*, R- and C-ratios are presented as color-coded images and values of δ are presented as orientation maps, where each line-bar shows the average collagen fiber orientation in the image plane.

2.4. Texture analysis of various SHG parameters

I_{SHG}, R- and C-ratios, δ, and the *DLP* images were subjected to image texture analysis where the relation between each pixel of the image and its nearest neighbors are considered. The texture analysis was performed using a gray-level co-occurrence matrix (GLCM). Fourteen textural parameters can be derived from the GLCM [30], amongst which some have previously been suggested as relevant to characterization of collagen SHG intensity images. The GLCM provides a second-order statistical representation of the distribution of gray levels within a specific ROI, which in turn provides the basis for textural analysis. GLCM is built by calculating the occurrence of a certain gray level pair i next to gray level j at the distance d along the direction γ. After

GLCM is obtained, the probability density function, $P_{d,\gamma}(i,j)$, of finding certain pairs of intensity i and j are calculated. Therefore, GLCM textural analysis considers the variation of pixel gray levels within a certain distance. Thereby, the forms, distributions and variation of the imaged objects, such as collagen fibers, can be tracked. Prior to textural analysis of the 8-bit images, the GLCM is calculated for each $\gamma = 0°$, $45°$, $90°$ and $135°$ angle directions, as well as each textural feature. The four obtained values are then averaged to account for possible variations regarding, for example, sample positioning on the imaging stage. These angles represent directions of neighbouring pixels and should not be mistaken for the PSG or PSA angles. A window size of 3×3 pixels used which then considers all 8 first neighboring pixels, and therefore accurately captures structures and variations associated to collagen fiber remodeling. Image processing and texture evaluations were then carried out using ImageJ [40] (post-image processing) and a custom-built toolkit incorporating MATLAB image processing toolbox scripts.

When studying collagen remodelling, it has been shown that some textural parameters are more informative than other with regards to the fibrous tissue morphology [35]. Performing a feature selection test based on Fisher score [41], only three GLCM features were found to be statistically significantly relevant in the context of defining structures in collagen fibers from PIPO SHG images obtained from lung biopsies. Here, three GLCM textural features- entropy, correlation and contrast- showed significant differences between tissue types, as judged by a Fischer score-based feature selection test. The three parameters are as follows.

The Entropy, S, is defined as:

$$S = -\sum_{i,j=0}^{N-1} P_{i,j} \log P_{i,j} \tag{4}$$

where N is the number of gray levels. Entropy measures the lack of spatial organization within the computational window and has been applied previously to collagen SHG intensity images [21,42,43]. High entropy means that the probability of finding certain paired gray levels are equal, i.e. it shows that the structures within the image are not organized, which corresponds to a rough texture.

The correlation parameter, ρ, quantifies a linear dependence of gray levels between two pixels separated by a certain distance d and it is defined as:

$$\rho = -\sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu)(j-\mu)}{\sigma^2} \right] \tag{5}$$

where μ is the mean and σ is the standard deviation of the gray levels. When an image shows low correlation, it means that the gray levels are generally independent from one another. When dealing with biological images, it can be associated with the fact that there is no regular structure within the image. However, if the correlation is high, then there is a high probability that one or several patterns repeat within the computational window. This parameter can be useful to quantify collagen fiber arrangement in normal and pathological tissues [35].

The third texture parameter, contrast or variability, ν, is defined as:

$$\nu = \sum_{i,j=0}^{N-1} (i-j)^2 P_{i,j} \tag{6}$$

and is a representation of pixels entirely similar to their neighbors, and it is very sensitive to large differences occurring inside the co-occurrence matrix.

The PIPO SHG microscopy images were used for the texture analysis. These images contain 81 polarization slices deduced from different combinations of PSG and PSA states, so that all slices were added to create a single polarization independent SHG intensity image for texture analysis. The texture analysis of the R- and C-ratios, δ and DLP was performed similarly.

2.5. Statistical analysis

The statistical parameters that were extracted from the *R*- and *C*-ratios as well as the δ and *DLP* distributions to investigate possible differences between collagen in normal and different stages of NSCLC, were the median and median absolute deviation (MAD). The latter is a measure of statistical dispersion and it is defined as the median of the absolute deviations from the median [44]. It has the advantage of being less sensitive to outliers than the standard deviation. To compare the textural features between collagen in normal and different tumor stages, the average of each feature was considered. After all statistical parameters from PIPO SHG measurements and texture analysis were extracted, D'Agostino-Pearson and Shapiro-Wilk normality test was used to assess the normality of the data. Linear mixed effect models were performed using GraphPad Prism version 8.1. Differences with p-value < 0.05 were considered as significantly different. The different ranges are indicated in each figure legend. To measure possible linear correlation between data sets a Pearson product-moment correlation coefficient was calculated.

3. Results

3.1. Tissue ultrastructure properties

The ultrastructure in each image pixel was examined with the linear PIPO SHG measurements. Examples of scans from normal lung parenchyma and the different tumor stages are shown in Fig. 1. It can be clearly seen that *R*-images in tumor samples are dominated by yellow and orange pixels compared to the control samples, which have mostly blue and green pixels (Fig. 1(c)). This can also be seen from the *R*-ratio distribution that shifts to higher values for tumor samples. The range of the color bar for the *C*-ratio (Fig. 1(d)) varies from blue to grey to red, depending on the out-of-image-plane collagen tilt angle, α, where blue indicates $\alpha<0$, grey represents $\alpha = 0$ and red shows fibres with $\alpha>0$. The *C*-ratio images show that, although the tissue samples were cut at random orientations, the majority of pixels contain collagen fibers that are oriented around $\alpha = 0$. The *C*-ratio image for stage-II and -III shows an alternating pattern of red, grey and blue, which is an indication of the collagen fiber waviness in 3D with the fiber orientation changes, respectively, from pointing into the image plane, being parallel to the image plane and pointing out of the image plane. None of the *C*-ratio distributions have a bimodal distribution, which was seen in a previous study in normal tendon where collagen fibers have an antiparallel arrangement [45].

Column (e) of Fig. 1 shows the fiber orientation (δ) map where each line (bar) represents the effective cylindrical axis of collagen fibres within each pixel. The histogram of collagen fibers orientation is also shown in the bottom left corner of Fig. 1(e) as a polar plot. A reduced waviness can be observed when comparing collagen fibers in tumor samples to the normal tissue. It can also be seen that the angle distributions in tumor samples are narrower than the normal samples, which is another indication of the reduction in the waviness.

The *DLP*, which is a measure of the average degree of linear polarization of the outgoing SHG signal from the sample, was also calculated for each pixel using Eqs.(3) and is presented as a color-coded image for each tissue in Fig. 1(f). The *DLP* distribution is closer to 1 for normal tissue compared to tumor samples. This can also be observed from the dominant red pixels in the normal tissue compared to the presence of some blue and green pixels in the tumor samples.

The statistical values extracted from the occurrence frequency histograms of *R*, *C*, δ and *DLP* for all scanned areas in normal and different tumor stages are presented in Fig. 2. No statistically significant differences observed between patients in control samples. Consequently all the control samples were pooled into one group. The *R*-ratio in tumor tissue of all stages is higher than in normal tissue, the differences being statistically significant in stage-II and -III (Fig. 2(a)). A trend can be seen between different tumor stages, where the *R*-ratio increases as the tumor progresses, but there were no statistically significant differences between stages. An increase in the *R*-ratio has been previously reported for lung cancer, where only stage-I was studied [17], as
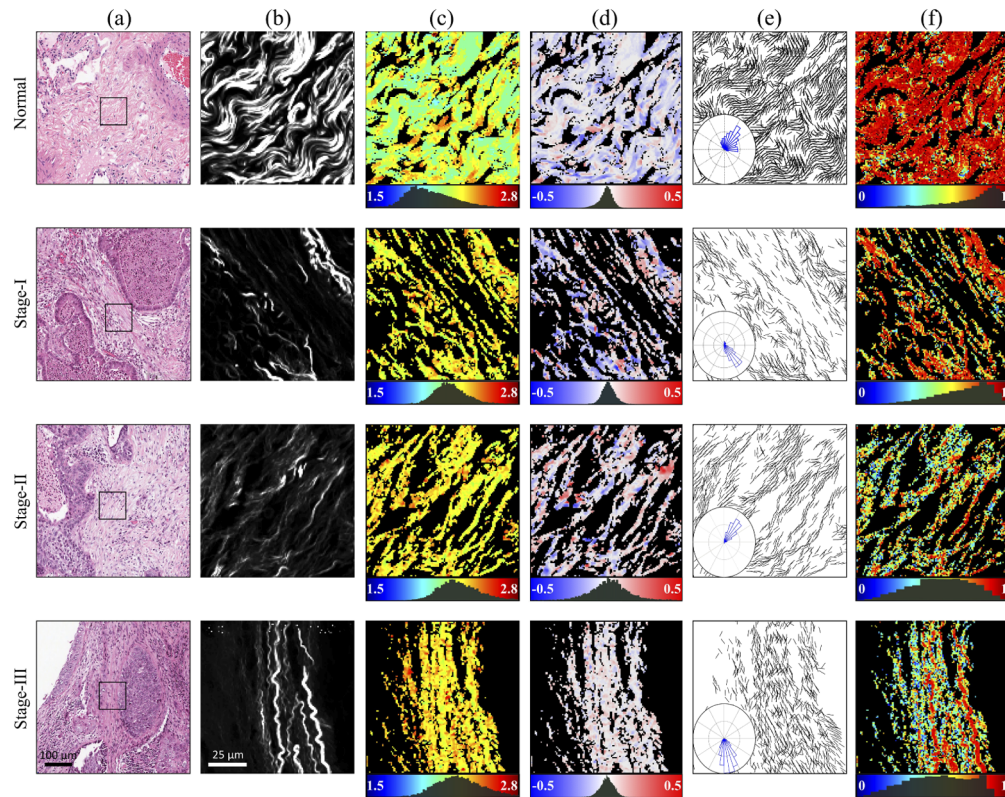
Fig. 1. PIPO SHG analysis of typical normal and NSCLC tissue sections. Column (a) shows bright-field microscopy images of the H&E stained tissues with square indicating the scanned area of 110 μm × 110 μm used for the SHG data analysis shown in columns (b) to (e). Column (b) visualizes the SHG intensity images (summation of all the polarization states) of the corresponding regions outlined by a black box in H&E images. Column (c) displays the *R*-map. Column (d) displays the *C*-ratio map. Column (e) shows the orientation map of the cylindrical axis, δ, for each fitted image pixel. The angle distribution is shown using a polar plot on the bottom left corner of the image. Column (f) displays the *DLP* map. The color bars also contain the occurrence pixel histogram of the corresponding images.

well as human breast [27], pancreas [28] and thyroid [26]. This increase in the *R*-ratio of tumor indicates less organized collagen fiber structure, and may be due to molecular changes of the collagen triple helices, different arrangement of triple helices forming fibers, and/or alteration in the organization of the fibers in the focal volume [17,24]. The SHG intensity and *R*-values for individual pixels did not correlate linearly (Pearson correlation coefficient of -0.12 for normal lung and -0.07, -0.09 and -0.08 for tumor stages-I, -II, and -III, respectively). The absence of linear correlation between the *R*-ratio and the SHG intensity indicates that there are additional factors, such as collagen concentration and fragmentation of fibrils, that influence the SHG intensity but are not reflected in the *R*-ratio to the same extent.
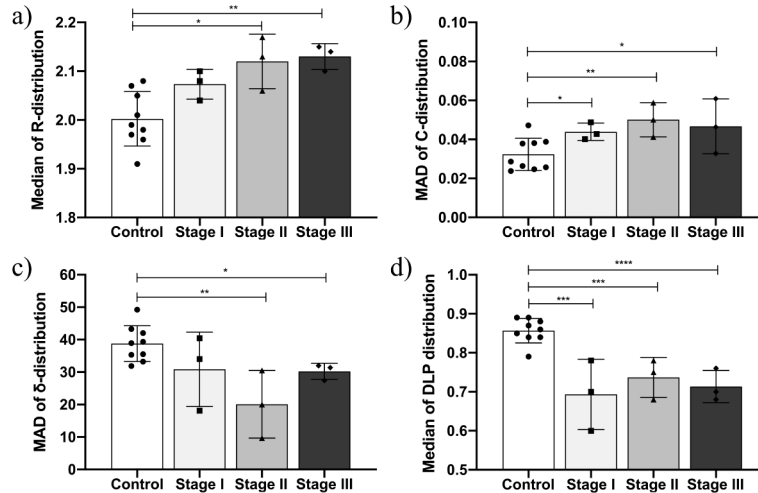

Fig. 2. Statistical parameters extracted from PIPO SHG analysis: (a) *R*-ratio median; (b) MAD of the *C*-ratio distribution, (c) MAD of δ; (d) Median of *DLP*. P-value <0.05*, <0.01**, 0.001***, <0.0001****. Error bars: ±1 standard deviation.

The use of the *C*-ratio distribution has been reported for differentiating between different collagen structures in heart tissue [46]. Here, since the *C*-ratio strongly depends on the out-of-plane orientation of the collagen fibers and the tissue samples were cut at a random orientation, the median absolute deviation (MAD) of *C*-ratio was used instead of the median. This shows that the *C*-ratio distribution is narrowest in normal lung tissue and that the difference with the tumor tissue is statistically significant at all 3 stages, indicating a larger spread of the out-of-plane orientation angle (α) that, in turn, indicates a more disorganized structure. It can also indicate alterations in the structure of collagen fibrils, which may influence the chiral χ_{xyz} or achiral χ_{zxx} component of the second-order susceptibility tensor of collagen. No significant differences were observed between the different tumor stages. As with the *R*-ratio, no correlation was found between the *C*-ratio and SHG intensity for each pixel (Pearson correlation coefficient= 0.01 for normal and -0.01, -0.02, and -0.05 for stage-I, -II, and -III, respectively).

Since the mean and the median of δ over the whole image depend on the sample orientation, they are not reported. Instead, the MAD of the δ values was extracted as a measure of the waviness of the collagen fibrils as well as the overall morphology of the sample. As Fig. 2(c) shows, the MAD values for tumor are smaller than for normal lung tissue, indicating that the collagen fibrils are straighter in tumor. This is also visible from Fig. 1(e), where the orientation lines become more aligned in the tumor samples. While the difference between normal and stage-I is not significant, there are statistically-significant differences between the normal tissue and both stage-II and stage-III. Both stage-I and -II follow a trend indicating that the collagen gets straighter as the stage of tumor increases, but this did not hold for stage-III NSCLC. This

can be because some selected ROIs are around the tumor lump, another curvature is introduced to the whole collagen in macro-scale which can cause an increase in the MAD of δ.

Figure 2(d) shows that the median *DLP* for all three stages are smaller than that for the normal tissue, the differences being statistically significant. However, no trend nor statistical difference was seen between the different tumor stages. Lower *DLP* means an increased depolarization of outgoing SHG due to higher ultrastructural disorder of collagen fibrils in the focal volume. It can also be due to fragmentation of collagen to smaller fiber segments within the focal volume. Hence, the SHG is emitted from uncorrelated domains and unpolarized scattering becomes significant compared to polarized SHG radiation. This feature was also observed in thyroid and breast tumors [26,27]. The Pearson correlation test showed very low correlation between the SHG intensity and the *DLP* (0.13 for normal and 0.26, 0.24, 0.27 for stage-I, -II and -III, respectively), although the correlation was still stronger than that between the *R*-ratio and the SHG intensity, suggesting that ultrastructural order leading to higher *DLP* has some influence on the SHG intensity.

3.2. *Texture analysis of different stages of NSCLC*

The texture analysis was conducted on the SHG intensity, I_{SHG}, as well as on all the *R*-, *C*-ratios, δ and *DLP* parameters extracted from the PIPO images. The I_{SHG}-image is obtained by summing all 81 images of the same area recorded at different combinations of incident laser and SHG signal polarizations, and so, it is polarization independent. Three textural features (entropy, correlation and contrast) in the SHG intensity images showed significant differences between tissue types. Figure 3(a) shows the average entropy for normal and all tumor stages. This is lowest for normal tissue and increases with stage, indicating decreasing uniformity and possibly increasing fragmentation. Such trend has been also observed in different diseases, such as asthma, and atherosclerosis [34,47], showing that collagen remodelling is present when there is incidence of fibrosis and/or scar formation.
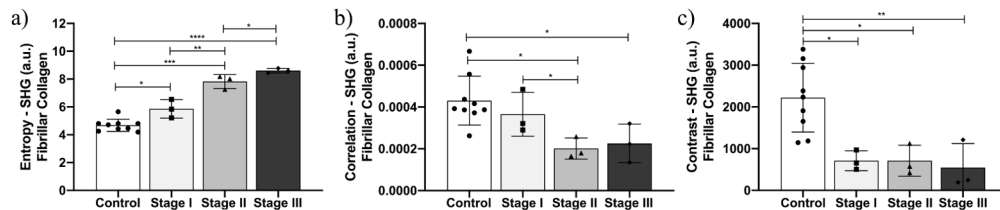

Fig. 3. Second-order parameters extracted from the GLCM: (a) Entropy, (b) Correlation, (c) Contrast. Features were individually calculated for each SHG image, summed across all the polarization states. P-value <0.05*, <0.01**, 0.001***, 0.0001****. Error bars: ±1standard deviation.

Figure 3(b) shows the correlation textural feature, with higher values for normal tissue than all tumor stages, with stages-II and -III being statistically significant. The higher correlation value for normal tissue indicates a more regular structural pattern, i.e. tumor tissues (especially stages-II and -III) have more isolated and fragmented collagen fibers, so that there is less probability of patterns of grey levels repeating within the computational window. Stages –II and –III were not statistically different from one another.

The contrast texture feature shown in Fig. 3(c) represents regions with varying SHG intensity in the neighboring pixels, higher values being associated with larger variability of gray levels within the computational window. Normal collagen shows a statistically-significant higher contrast compared to all three NSCLC stages. The contrast decreases as a function of tumor stage, however there are no significant differences between stages. Higher contrast indicates that normal fibers have better order, i.e. are less fragmented, leading to brighter SHG than the surroundings.

The entropy, correlation and contrast textural features for images of *R*- and *C*-ratios, δ and *DLP* did not show statistically significant differences but some trends are apparent. Firstly, the *R*-ratio entropy decreases in stage-I and stage-II tumor compared to normal tissue, while the correlation and contrast exhibit no differences between normal and tumor samples. From the ultrastructure analysis we know that mean of the *R*-ratio shifts to the higher values and according to the texture analysis result this increase occur consistently throughout the image. Secondly, in the δ-value texture analysis the correlation parameter shows a decreasing trend for tumor compared to normal tissue. Since correlation quantifies the dependence of two pixels separated by distance *d*, lower correlation means that the pixels are generally independent from one another. This can be explained by high spatial-frequency waviness of the fibers in the tumor. A high spatial frequency waviness of collagen occurs on a micron-scale, while the whole fibre on a few tens of microns scale get straighter as explained by the MAD of δ. On the other hand, the fibers in normal tissue follow a low waviness on a micron-scale while having a wavy structure on a larger scale of tenths of microns. Lastly, the contrast parameter in the *DLP* texture analysis shows higher values for normal than all stages of tumor tissues, indicating a better order in the normal collagen fibers. This is similar to the contrast parameter of the SHG intensity texture analysis. The fact that there is higher correlation between *DLP* and SHG intensity than between the *R*-ratio and SHG intensity indicates that *DLP* is more sensitive to collagen disorder.

4. Discussion

We have used polarimetric SHG microscopy to study the impact of NSCLC tumor stage on the collagen structure. The ultrastructure analysis provides *R*, *C*, δ and *DLP* parameters for each image pixel, yielding information about the ultrastructural organization of collagen fibers within pixels (through *R*, *C*, and *DLP*) and about the effective collagen orientation (through δ). In addition, texture analysis of these parameters and of the SHG intensity were used to investigate the microstructural variation of collagen fibers between the nearest neighboring pixels in the computational window that is probed across the scanned area. Three textural parameters were extracted: entropy, correlation and contrast.

The *R*-ratio distribution showed lower values in normal tissue and increase with tumor stage. One possible mechanism is the re-organization of collagen fiber matrices. This is due to the overexpression of ECM proteinases [48] and increase in collagen fiber stiffness, demonstrating that the ECM organization and stiffness is a key regulator of tumor progression. A correlation between integrin α11 expression and collagen stiffness has been reported [15] and expression of collagen-binding integrins is known to remodel collagen type-I by mediating contraction of the collagen lattices [15]. Images obtained from atomic force microscopy have revealed that collagen fibers are less organized and the Young's modulus varies significantly as a function of integrin α11 expression [15]. These observations could be linked to ultrastructural properties of collagen which impacts the *R*-ratio values.

The δ-distributions showed that collagen in tumor tissue has increased fiber alignment compared to the normal tissue. The lower values of the contrast parameter of the SHG intensity can also be interpreted as a better aligned and oriented structure. This behavior has been shown through another study on NSCLC [15] as well as other types of carcinoma, including breast [27,49–51] and pancreas [16]. Increase in the collagen fiber alignment and orientation is one of the proposed mechanisms of metastasic spread from solid tumors [52].

The MAD of the *C*-ratio distributions trended towards increase in tumor stage. Since the *C*-ratio is a function of $\sin\alpha$, this may indicate that the spread of collagen fibers out of the image plane is increasing, in contrast to what is observed for δ. The random-orientation sectioning of the tissue should eliminate the differences between the in-plane and out-of-plane angle distributions. This points to increase in the molecular χ_{xyz}/χ_{zxx} distribution in tumor, possibly due to collagen fragmentation.

The *DLP* values were smaller in malignant than normal lung tissue, indicating increased disorder in collagen fibers of tumor tissue. Collagen fragmentation is consistent with increase in the entropy in the SHG intensity texture analysis, which is lower for normal tissue and increase with tumor stage, indicating loss of structural organization. The correlation parameter from texture analysis of the SHG intensity also suggests that the collagen fibers in the computational window are more fragmented for malignant tissue. It has been previously shown that integrin $\alpha 11$ can influence the linearization of collagen fibers adjacent to the tumor cells [15] and that disorganized collagen fibers at the tumor cell/stroma interface can be created as the result of direct contact between tumor cells and cancer associated fibroblasts (CAF) [52].

5. Conclusions

The combination of SHG polarimetric analysis and texture analysis reveals significant differences in the collagen structure between NLSC and normal lung tissue in the *R*, *C*, δ, and *DLP* parameters from polarimetric measurements, and the entropy, correlation and contrast from the SHG intensity texture analysis. With the tumor progression, the collagen structure becomes more disordered and fragmented on the submicron level, while at the same time remodeled fibers become straighter and preferentially microstructurally aligned. Hence, polarimetric SHG microscopy in combination with texture analysis can provide a powerful tool to investigate the extracellular tumor microenvironment. The extension of this study to a large tissue area investigation and combination with other histopathology techniques may be of value for more accurate cancer diagnostics and staging.

Funding

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. D. C. Paech, A. R. Weston, N. Pavlakis, A. Gill, N. Rajan, H. Barraclough, B. Fitzgerald, and M. Van Kooten, "A systematic review of the interobserver variability for histology in the differentiation between squamous and nonsquamous non-small cell lung cancer," J. Thorac. Oncol. **6**(1), 55–63 (2011).
2. World Health Organization, "Cancer," http://www.who.int/mediacentre/factsheets/fs297/en/.
3. M. Egeblad, A. J. Ewald, H. a. Askautrud, M. L. Truitt, B. E. Welm, E. Bainbridge, G. Peeters, M. F. Krummel, and Z. Werb, "Visualizing stromal cell dynamics in different tumor microenvironments by spinning disk confocal microscopy," Dis. Models Mech. **1**(2-3), 155–167 (2008).
4. K. R. Levental, H. Yu, L. Kass, J. N. Lakins, M. Egeblad, J. T. Erler, S. F. T. Fong, K. Csiszar, A. Giaccia, W. Weninger, M. Yamauchi, D. L. Gasser, and V. M. Weaver, "Matrix Crosslinking Forces Tumor Progression by Enhancing Integrin Signaling," Cell **139**(5), 891–906 (2009).
5. E. Y. Lin, J. F. Li, L. Gnatovskiy, Y. Deng, L. Zhu, D. A. Grzesik, H. Qian, X. N. Xue, and J. W. Pollard, "Macrophages regulate the angiogenic switch in a mouse model of breast cancer," Cancer Res. **66**(23), 11238–11246 (2006).
6. Z. Chen, C. M. Fillmore, P. S. Hammerman, C. F. Kim, and K.-K. Wong, "Non-small-cell lung cancers: a heterogeneous set of diseases," Nat. Rev. Cancer **14**(8), 535–546 (2014).
7. C. Bonnans, J. Chou, and Z. Werb, "Remodelling the extracellular matrix in development and disease," Nat. Rev. Mol. Cell Biol. **15**(12), 786–801 (2014).
8. S. M. Pupa, S. Ménard, S. Forti, and E. Tagliabue, "New insights into the role of extracellular matrix during tumor onset and progression," J. Cell. Physiol. **192**(3), 259–267 (2002).
9. N. Théret, O. Musso, B. Turlin, D. Lotrian, P. Bioulac-Sage, J. P. Campion, K. Boudjéma, and B. Clément, "Increased extracellular matrix remodeling is associated with tumor progression in human hepatocellular carcinomas," Hepatology **34**(1), 82–88 (2001).
10. A. Cho, V. M. Howell, and E. K. Colvin, "The Extracellular Matrix in Epithelial Ovarian Cancer – A Piece of a Puzzle," Front. Oncol. **5**, 1–16 (2015).

11. R. Kalluri, "Basement membranes: Structure, assembly and role in tumour angiogenesis," Nat. Rev. Cancer **3**(6), 422–433 (2003).

12. K. Burke, P. Tang, and E. Brown, "Second harmonic generation reveals matrix alterations during breast tumor progression," J. Biomed. Opt. **18**(3), 031106 (2012).

13. P. P. Provenzano, K. W. Eliceiri, J. M. Campbell, D. R. Inman, J. G. White, and P. J. Keely, "Collagen reorganization at the tumor-stromal interface facilitates local invasion," BMC Med. **4**(1), 38 (2006).

14. P. P. Provenzano, D. R. Inman, K. W. Eliceiri, J. G. Knittel, L. Yan, C. T. Rueden, J. G. White, and P. J. Keely, "Collagen density promotes mammary tumor initiation and progression," BMC Med. **6**(1), 11 (2008).

15. R. Navab, D. Strumpf, C. To, E. Pasko, K. S. Kim, C. J. Park, J. Hai, J. Liu, J. Jonkman, M. Barczyk, B. Bandarchi, Y. H. Wang, K. Venkat, E. Ibrahimov, N. A. Pham, C. Ng, N. Radulovich, C.-Q. Zhu, M. Pintilie, D. Wang, A. Lu, I. Jurisica, G. C. Walker, D. Gullberg, and M.-S. Tsao, "Integrin $\alpha 11 \beta 1$ regulates cancer stromal stiffness and promotes tumorigenicity and metastasis in non-small cell lung cancer," Oncogene **35**(15), 1899–1908 (2016).

16. C. R. Drifka, A. G. Loeffler, K. Mathewson, A. Keikhosravi, J. C. Eickhoff, Y. Liu, S. M. Weber, K. W. Eliceiri, and W. J. Kao, "Highly aligned stromal collagen is a negative prognostic factor following pancreatic ductal adenocarcinoma resection," Oncotarget **7**(46), 76197–76213 (2016).

17. A. Golaraei, R. Cisek, S. Krouglov, R. Navab, C. Niu, S. Sakashita, K. Yasufuku, M. S. Tsao, B. C. Wilson, and V. Barzda, "Characterization of collagen in non-small cell lung carcinoma with second harmonic polarization microscopy," Biomed. Opt. Express **5**(10), 3562–3567 (2014).

18. S. Roth and I. Freund, "Second harmonic generation in collagen," J. Chem. Phys. **70**(4), 1637–1643 (1979).

19. I. Freund and M. Deutsch, "Second-harmonic microscopy of biological tissue," Opt. Lett. **11**(2), 94–96 (1986).

20. I. Freund, M. Deutsch, and A. Sprecher, "Connective tissue polarity. Optical second-harmonic microscopy, crossed-beam summation, and small-angle scattering in rat-tail tendon," Biophys. J. **50**(4), 693–712 (1986).

21. P. J. Campagnola and C. Y. Dong, "Second harmonic generation microscopy: Principles and applications to disease diagnosis," Laser Photonics Rev. **5**(1), 13–26 (2011).

22. P. J. Campagnola, A. C. Millard, M. Terasaki, P. E. Hoppe, C. J. Malone, and W. A. Mohler, "Three-dimensional high-resolution second-harmonic generation imaging of endogenous structural proteins in biological tissues," Biophys. J. **82**(1), 493–508 (2002).

23. A. E. Tuer, S. Krouglov, N. Prent, R. Cisek, D. Sandkuijl, K. Yasufuku, B. C. Wilson, and V. Barzda, "Nonlinear optical properties of type I collagen fibers studied by polarization dependent second harmonic generation microscopy," J. Phys. Chem. B **115**(44), 12759–12769 (2011).

24. A. E. Tuer, M. K. Akens, S. Krouglov, D. Sandkuijl, B. C. Wilson, C. M. Whyne, and V. Barzda, "Hierarchical model of fibrillar collagen organization for interpreting the second-order susceptibility tensors in biological tissue," Biophys. J. **103**(10), 2093–2105 (2012).

25. A. Golaraei, K. Kontenis, K. Mirsanaye, S. Krouglov, M. K. Akens, B. C. Wilson, and V. Barzda, "Complex Susceptibilities and Chiroptical Effects of Collagen Measured with Polarimetric Second-Harmonic Generation Microscopy," Sci. Rep. **9**(1), 12488 (2019).

26. D. Tokarz, R. Cisek, A. Golaraei, S. L. Asa, V. Barzda, and B. C. Wilson, "Ultrastructural features of collagen in thyroid carcinoma tissue observed by polarization second harmonic generation microscopy," Biomed. Opt. Express **6**(9), 3475–3481 (2015).

27. A. Golaraei, L. Kontenis, R. Cisek, D. Tokarz, S. J. Done, B. C. Wilson, and V. Barzda, "Changes of collagen ultrastructure in breast cancer tissue determined by second-harmonic generation double Stokes-Mueller polarimetric microscopy," Biomed. Opt. Express **7**(10), 4054–4068 (2016).

28. D. Tokarz, R. Cisek, A. Joseph, A. Golaraei, K. Mirsanaye, S. Krouglov, S. L. Asa, B. C. Wilson, and V. Barzda, "Characterization of pancreatic cancer tissue using multiphoton excitation fluorescence and polarization-sensitive harmonic generation microscopy," Front. Oncol. **9**, 272 (2019).

29. M. Bevk and I. Kononenko, "A statistical approach to texture description of medical images: A preliminary study," *Proc. IEEE Symp. on Comput. Med. Syst.* pp. 239–244 (2002).

30. R. M. Haralick and K. Shanmugam, "Textural Features for Image Classification," IEEE Transactions on Syst. Man, Cybern. **SMC-3**(6), 610–621 (1973).

31. M. Strupler, M. Hernest, C. Fligny, J.-L. Martin, P.-L. Tharaux, and M.-C. Schanne-Klein, "Second harmonic microscopy to quantify renal interstitial fibrosis and arterial remodeling," J. Biomed. Opt. **13**(5), 054041 (2008).

32. A.-M. Pena, A. Fabre, D. Débarre, J. Marchal-Somme, B. Crestani, J.-L. Martin, E. Beaurepaire, and M.-C. Schanne-Klein, "Three-dimensional investigation and scoring of extracellular matrix remodeling during lung fibrosis using multiphoton microscopy," Microsc. Res. Tech. **70**(2), 162–170 (2007).

33. L. Gailhouste, Y. L. Grand, C. Odin, D. Guyader, B. Turlin, F. Ezan, Y. Désille, T. Guilbert, A. Bessard, C. Frémin, N. Theret, and G. Baffet, "Fibrillar collagen scoring by second harmonic microscopy: A new tool in the assessment of liver fibrosis," J. Hepatol. **52**(3), 398–406 (2010).

34. L. B. Mostaço-Guidolin, A. C-T Ko, D. P. Popescu, M. S. D Smith, E. K. Kohlenberg, M. Shiomi, A. Major, and M. G. Sowa, "Evaluation of texture parameters for the quantitative description of multimodal nonlinear optical images from atherosclerotic rabbit arteries," Phys. Med. Biol. **56**(16), 5319–5334 (2011).

35. L. B. Mostaço-Guidolin, A. C.-T. Ko, F. Wang, B. Xiang, M. Hewko, G. Tian, A. Major, M. Shiomi, and M. G. Sowa, "Collagen morphology and texture analysis: from statistics to classification," Sci. Rep. **3**(1), 2190 (2013).

36. L. B. Mostaço-Guidolin, E. T. Osei, S. Hajimohammadi, J. Ullah, and T. Hackett, "Novel Non-Linear Optical Imaging To Understand The Composition Of Fibrilar Collagen And Elastin In Remodeled Asthmatic Airways," Am. J. Respir. Crit. Care Med. **2016**, 6173 (2016).

37. A. Major, R. Cisek, and V. Barzda, "Femtosecond Yb:KGd(WO(4))(2) laser oscillator pumped by a high power fiber-coupled diode laser module," Opt. Express **14**(25), 12163–12168 (2006).

38. C. A. Dailey, B. J. Burke, and G. J. Simpson, "The general failure of Kleinman symmetry in practical nonlinear optical applications," Chem. Phys. Lett. **390**(1-3), 8–13 (2004).

39. T. F. Coleman and Y. Li, "An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds," (1996).

40. C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," Nat. Methods **9**(7), 671–675 (2012).

41. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," J. Mach. Learn. Res. **3**, 1157–1182 (2003).

42. P. J. Campagnola and L. M. Loew, "Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms," Nat. Biotechnol. **21**(11), 1356–1360 (2003).

43. L. Mostaço-Guidolin, N. Rosin, and T.-L. Hackett, "Imaging Collagen in Scar Tissue: Developments in Second Harmonic Generation Microscopy for Biomedical Applications," Int. J. Mol. Sci. **18**(8), 1772 (2017).

44. C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," J. Exp. Soc. Psychol. **49**(4), 764–766 (2013).

45. A. Golaraei, K. Mirsanaye, Y. Ro, S. Krouglov, M. K. Akens, B. C. Wilson, and V. Barzda, "Collagen chirality and three-dimensional orientation studied with polarimetric second-harmonic generation microscopy," J. Biophotonics **12**(1), e201800241 (2019).

46. K. Mirsanaye, A. Golaraei, F. Habach, E. Žurauskas, J. Venius, R. Rotomskis, and V. Barzda, "Polar organization of collagen in human cardiac tissue revealed with polarimetric second-harmonic generation microscopy," Biomed. Opt. Express **10**(10), 5025–5030 (2019).

47. L. B. Mostaço-Guidolin, E. T. Osei, J. Ullah, S. Hajimohammadi, M. Fouadi, X. Li, V. Li, F. Shaheen, C. X. Yang, F. Chu, D. J. Cole, C. A. Brandsma, I. H. Heijink, G. N. Maksym, D. Walker, and T. L. Hackett, "Defective fibrillar collagen organization by fibroblasts contributes to airway remodeling in asthma," Am. J. Respir. Crit. Care Med. **200**(4), 431–443 (2019).

48. M. Hidalgo and S. G. Eckhardt, "Development of matrix metalloproteinase inhibitors in cancer therapy," J. Natl. Cancer Inst. **93**(3), 178–193 (2001).

49. A. Brabrand, I. I. Kariuki, M. J. Engstrøm, O. A. Haugen, L. A. Dyrnes, B. O. Åsvold, M. B. Lilledahl, and A. M. Bofin, "Alterations in collagen fibre patterns in breast cancer. A premise for tumour invasiveness?" APMIS **123**(1), 1–8 (2015).

50. M. W. Conklin, J. C. Eickhoff, K. M. Riching, C. a. Pehlke, K. W. Eliceiri, P. P. Provenzano, A. Friedl, and P. J. Keely, "Aligned collagen is a prognostic signature for survival in human breast carcinoma," Am. J. Pathol. **178**(3), 1221–1232 (2011).

51. W. Han, S. Chen, W. Yuan, Q. Fan, J. Tian, X. Wang, L. Chen, X. Zhang, W. Wei, R. Liu, J. Qu, Y. Jiao, R. H. Austin, and L. Liu, "Oriented collagen fibers direct tumor cell intravasation," Proc. Natl. Acad. Sci. U. S. A. **113**(40), 11208–11213 (2016).

52. H. H. Popper, "Progression and metastasis of lung cancer," Cancer Metastasis Rev. **35**(1), 75–91 (2016).